THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 62, No. 8, October 1983
Printed in U.S.A.

Upper Bounds on the Minimum Distance of Trellis Codes

By A. R. CALDERBANK,* J. E. MAZO,* and H. M. SHAPIRO†

(Manuscript received March 16, 1983)

A trellis code is a "sliding window" metbod of encoding a binary data stream into a sequence of real numbers that are input to a noisy transmission channel. When a trellis code is used to encode data at tbe rate of k bits/channel symbol, each channel input will depend not only on the most recent block of k data hits to enter the encoder hut will also depend on, say, the ν bits preceding this block. The ν bits determine the state of the encoder and the most recent block of k hits generates the channel symhol conditional on tbe encoder state. The performance of trellis codes, like that of hlock codes, dependa on a suitably defined minimum-distance property of the code. In this paper we obtain upper hounds on this minimum distance that are simple functions of k and ν. These results also provide a lower hound on the numher of states required to achieve a specific coding gain.

I. INTRODUCTION

In this paper we are concerned with transmission of digital data using trellis codes to gain some noise immunity over standard uncoded methods. We assume pulse amplitude modulation whereby the values of the transmitted data are estimated from a sequence of samples r^j generated by a receiver. These output samples are often modeled as

$$r^j = x^j + n^j, \tag{1}$$

where x^j is a real number sequence determined by the source sequence of binary data and n^j is an independent zero-mean white Gaussian

* Bell Laboratories. † Swarthmore College, Swarthmore, Pennsylvania.

noise sequence of variance σ^2. For uncoded transmission at rate k bits/symbol, x^j takes on one of 2^k fixed values. Error performance may be improved using coding, but if we insist on transmitting at rate k bits/symbol then we must increase the number of possible values taken by the x^j. We can choose either a block or tree (trellis) structure for the code. In this paper we consider only trellis codes. The performance of trellis codes, like that of block codes, depends on a suitably defined minimum-distance property of the code. We obtain upper bounds on this minimum distance, $d_{\min}$. The analogous problem for block codes is well studied, but little work has been done on distance properties of trellis codes.[1,2]

We assume the following model for encoding the binary data (i.e., choosing the x^j) prior to transmission over the Gaussian channel. Regard the incoming binary digits as partitioned into blocks of k consecutive bits. The real number x^j is to be a time-independent function of the most recent k-bit block and also of the ν bits preceding this block. Thus if $\{a_i\}$ is the binary data sequence, we assume

$$x^j = x(a_{jk}, a_{jk-1}, \cdots, a_{jk-(k-1)}; a_{(j-1)k}, \cdots, a_{(j-1)k-(\nu-1)}). \tag{2}$$

This is an example of a k-bit/symbol trellis code. We regard the ν "old" bits as determining the state of the encoder (there are 2^ν possible states) and the k "new" bits as generating the channel symbol (there are 2^k possible symbols) conditional on the encoder state. The trellis structure is made evident by drawing an example. Fig. 1 shows the case $k = 1$, $\nu = 2$.

If, in this example, the encoder is in state (00) at time j, and the next bit (block of $k = 1$ bits) to be transmitted is a 1, then we transmit the symbol $x(100)$ and move to state (10).

Other trellis codes exist. For example, we could define a code with just three trellis states or the symbols x^j could also depend on the time index j. However, we shall only consider trellis codes determined by (2). The trellis structure of (2) is identical to that of linear algebraic

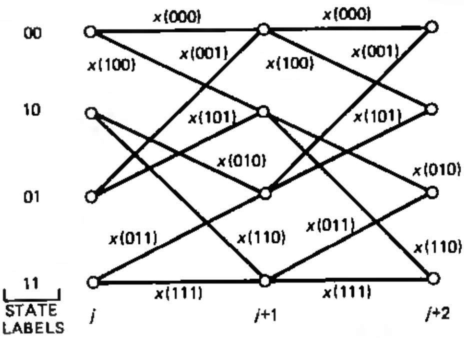

Fig. 1—Diagram of a trellis code.

convolutional codes. We use tbe term sliding window trellis codes for trellis codes determined by (2).

To simplify the discussion in tbe text we shall assume tbat k divides ν. The general case is treated in Appendix B.

The problem we consider involves certain distance properties of trellis codes. To motivate it, consider the decoding problem. Optimum decoding involves finding the most likely path through the trellis, given the observed sequence (1).[3] Typically, tbe patb cbosen will not coincide with the correct path for all time but will occasionally diverge from it and remerge at a later time. This is called an error event, and we generically denote it by the letter E. For example, with the trellis in Fig. 1, $x(000)$ may bave been sent several times in succession, resulting in the straight path sbown in Fig. 2, but noise may have caused the decoder to choose an alternate patb. In Fig. 2 tbe decoder chose the symbols $x(100)$, $x(010)$, $x(001)$ instead of $x(000)$, $x(000)$, $x(000)$.

An error event E of length L lasts from time i to time $i + L$, the decoder having decided upon the symbol sequence $\hat{x}^{i+1}, \cdots, \hat{x}^{i+L}$ instead of the correct sequence $x^{i+1}, \cdots, x^{i+L}$. The (squared) Euclidean distance d^2 $(= d^2(E))$ between the two paths of E is given by

$$d^2 = \sum_{j=i+1}^{i+L} (x^j - \hat{x}^j)^2 \tag{3}$$

and is crucial to determining the probability $P(E)$ of an error event E. With the white noise assumption made in (1), $P(E)$ is easy to calculate and, when $d^2 \gg \sigma^2$, it is approximately given by

$$P(E) \approx \exp\left(-\frac{d^2}{8\sigma^2}\right). \tag{4}$$

Equation (4) leads us to expect that, for small noise, symbol error probabilities will be determined by error events baving the smallest minimum distance between their two patbs and it becomes of interest to design codes that have good minimum-distance properties in this sense. Such designs have recently been considered by Ungerboeck, who obtained on the order of 3-dB performance improvements (factor

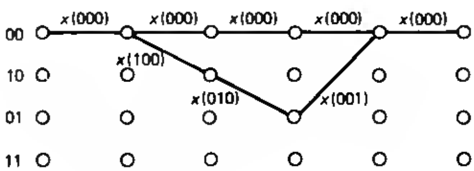

Fig. 2—Example of an error event.

of 2 in minimum distance) over the uncoded case for $k = 1, 2$, and four or eight states in the trellis.[4]

Ungerboeck based his designs on a computer search of binary convolutional codes with 2^ν states, rate $k/(k+1)$, and a particular mapping of the output binary $(k+1)$ tuples to 2^{k+1} equally spaced channel symhols (± 1, ± 3, etc.). His use of convolutional codes thus conforms to the general scheme of (2), which implies the same trellis structure as described herein. However, his a priori choice of only 2^{k+1} equally spaced channel symhols is certainly restrictive in principle. In this paper we consider the natural question of how large $d_{\min}^2/P$ can be made if these restrictions are removed. Here, $d_{\min}$ is the minimum distance hetween all pairs of paths associated with error events in the trellis, and P is the average transmitted power.

Section II gives a detailed description of the trellis structure and of error events. If S is a finite set of error events, then

$$\min_{E \in S} \{d^2(E)\} \leq \frac{1}{|S|} \sum_{E \in S} d^2(E), \tag{5}$$

since the minimum of a set of real numbers is bounded above by their average. This observation is the basis of our first two bounds. The first and simplest bound is

$$\frac{d_{\min}^2}{P} \leq 4\left(1 + \frac{\nu}{k}\right), \tag{6}$$

which is obtained in Section III. A more detailed analysis in Section IV gives

$$\frac{d_{\min}^2}{P} \leq \frac{2^{k+1}}{2^k - 1}\left(1 + \frac{\nu}{k}\right), \tag{7}$$

which is stronger than (6) provided $k > 1$. Let T be another finite set of error events and let $r_1, r_2 \geq 0$ he real numhers satisfying $r_1 + r_2 = 1$. Then,

$$\min_{E \in S \cup T} \{d^2(E)\} \leq r_1 \left(\frac{1}{|S|} \sum_{E \in S} d^2(E)\right) + r_2 \left(\frac{1}{|T|} \sum_{E \in T} d^2(E)\right), \tag{8}$$

since the minimum of a set of real numbers is hounded above by any weighted average of those numbers. In Section V, by choosing S, T, r_1, and r_2, appropriately, we prove

$$\frac{d_{\min}^2}{P} \leq \left(\frac{2^{2k+1}}{2^{2k} - 1}\right)\left(2 + \frac{\nu}{k}\right). \tag{9}$$

This bound is stronger than (7) provided $\nu > k(2^k - 1)$. Combining (7)

and (9) we have

$$\frac{d_{\min}^2}{P} \leqslant \min\left[\frac{2^{k+1}}{2^k-1}\left(1+\frac{\nu}{k}\right), \frac{2^{2k+1}}{2^{2k}-1}\left(2+\frac{\nu}{k}\right)\right]. \tag{10}$$

Extensions of bounds (6), (7), and (9) to tbe case wben k does not divide ν are given in Appendix B.

II. A GROUP ACTION ON THE TRELLIS

In later sections we obtain upper bounds on $d_{\min}^2/P$ by considering sets of error events that are fixed by a group of symmetries of the trellis. In this section we describe the group.

We consider trellis codes with 2^ν states transmitting k bits/cbannel symbol and for simplicity we assume tbat k divides ν. States are labelled with binary ν tuples, and edges of tbe trellis are labelled with binary $\nu + k$ tuples. We identify tbe binary r tuple $(b_0, \cdots, b_{r-1})$ with the integer

$$b_0 2^0 + b_1 2^1 + \cdots + b_{r-1} 2^{r-1}.$$

The states are labelled witb binary ν tuples $00 \cdots 0$, $10 \cdots 0$, $010 \cdots 0$, $110 \cdots 0$, $\cdots$, $11 \cdots 1$, in increasing order, from top to bottom as in Fig. 1. The edges are labelled with binary $\nu + k$ tuples $x_0 = x(0 \cdots 0)$, $x_1 = x(10 \cdots 0)$, $x_2 = x(010 \cdots 0)$, $x_3 = x(110 \cdots 0)$, $\cdots$, $x_{2^{\nu+k}-1} = x(11 \cdots 1)$, also in increasing order, from top to bottom as in Fig. 1. Set $N = (k + \nu)/k$. If we write an edge label as $x(s_0, \cdots, s_{N-1})$, then it will be understood that each s_j is a binary k tuple. A "+" appearing in the argument of a label means bit-by-bit modulo 2 addition. A similar notation will be used for states.

We define a group of symmetries of the trellis. These symmetries will map error events of lengtb L to error events of lengtb L. For eacb binary $\nu + k$ tuple t, we define a permutation g_t of the edge labels $x(s)$ by the rule

$$g_t(x(s)) = x(s + t). \tag{11}$$

For example, when $k = 1$, $\nu = 2$, and $t = (010)$,

$$\underline{x} = \begin{bmatrix} x(000) \\ x(100) \\ x(010) \\ x(110) \\ x(001) \\ x(101) \\ x(011) \\ x(111) \end{bmatrix} \mapsto g_{010}(\underline{x}) = \begin{bmatrix} x(010) \\ x(110) \\ x(000) \\ x(100) \\ x(011) \\ x(111) \\ x(001) \\ x(101) \end{bmatrix}. \tag{12}$$

This may also be written

$$g_{010}(\underset{\sim}{x}) = T\underset{\sim}{x}, \tag{13}$$

where T is tbe permutation matrix

$$T = \begin{bmatrix} 0 & 0 & 1 & 0 & & & & \\ 0 & 0 & 0 & 1 & & & & \\ 1 & 0 & 0 & 0 & & & & \\ 0 & 1 & 0 & 0 & & & & \\ & & & & 0 & 0 & 1 & 0 \\ & & & & 0 & 0 & 0 & 1 \\ & & & & 1 & 0 & 0 & 0 \\ & & & & 0 & 1 & 0 & 0 \end{bmatrix}. \tag{14}$$

If $\mathbf{G}_{k,\nu} = \{g_t \,|\, t \text{ is a binary } \nu + k \text{ tuple}\}$, then $\mathbf{G}_{k,\nu}$ is an abelian group of order $2^{k+\nu}$, and every element g_t of $\mathbf{G}_{k,\nu}$ satisfies $g_t^2 = e$, where e is the group identity.

Lemma 1: Any pair of edge labels is interchanged by a unique group element.

Proof: Edge labels $x(s)$ and $x(u)$ are interchanged only by g_{s+u}. □

We call the time sections (0, 1), (1, 2), $\cdots$ the components of the trellis. We shall now show how to choose binary $\nu + k$ tuples $t = t^0, t^1, \cdots$ so that if g_{t^i} is applied to the edges in component i, then an error event of length L is always mapped to another error event of length L. It is, in general, necessary to choose a different g_t for each component since if we simply apply the same permutation g_t to the edges in every component, tben an error event E need not be transformed to another error event. Thus, if g_{010} is applied to each component of the error event shown in Fig. 2, then we obtain the edges sbown in Fig. 3. The permutation g_{010} transforms the edge labelled $x(uvw)$, joining state vw and state uv, into the edge labelled $x(u(1 + v)w)$, joining state $(1 + v)w$ and state $u(1 + v)$. If $t = t^0 = 010$, then g_{010} permutes the encoder states at time 0 by the rule

$$vw \mapsto (1 + v)w, \tag{15}$$

and permutes the encoder states at time 1 by tbe rule

$$uv \mapsto u(1 + v). \tag{16}$$

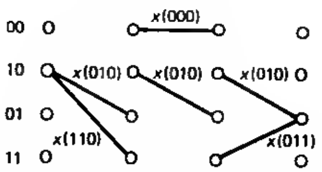

Fig. 3—Permutation g_{010} applied to all edges of an error event.

Similarly, the permutation g_{t^1} permutes encoder states at time 1 and encoder states at time 2. If we want to map error events to error events, then the action of g_{t^1} on encoder states at time 1 must be given by (16). Choose $t^1 = 001$, $t^2 = 100$, $t^3 = 010$, $t^4 = 001$, $\cdots$. The action of g_{t^0}, g_{t^1}, and g_{t^2} on components 0, 1, and 2 is shown in Fig. 4. Thus the sequence $(g_{t^0}, g_{t^1}, g_{t^2}, \cdots)$ transforms the error event shown in Fig. 2 to the error event shown in Fig. 5.

For general k and ν, let $t = t^0 = (t_0, \cdots, t_{N-1})$ where $N = (k + \nu)/k$ and $t_0, \cdots, t_{N-1}$ are binary k tuples. Let $t^1 = (t_{N-1}, t_0, \cdots, t_{N-2})$ be the vector obtained from t^0 by cycling the blocks of k bits to the right and moving the last block, t_{N-1}, to the front. Repeat this operation i times to obtain $t^i = (t_{N-i}, \cdots, t_{N-1}, t_0, \cdots, t_{N-i-1})$. For $i \geqslant N$ we view i as an integer modulo N. Thus $t^N = t^0 = t$, $t^{N+1} = t^1$, $\cdots$. The action of g_{t^i} on encoder states at time i coincides with that of $g_{t^{i-1}}$ being given by the rule

$$s \mapsto (t_{N-i+1}, \cdots, t_{N-1}, t_0, \cdots t_{N-i-1}) + s. \tag{17}$$

If $\mathbf{G}^*_{k,\nu} = \{(g_{t^0}, g_{t^1}, \cdots) \mid t^0 \text{ is a binary } k \text{ tuple}\}$, then $\mathbf{G}^*_{k,\nu}$ is a group of $2^{\nu+k}$ symmetries of the trellis. The group $\mathbf{G}^*_{k,\nu}$ is abelian, and every element has order 2. We denote $(g_{t^0}, g_{t^1}, \cdots)$ by $g^*_{t^0}$, since it is determined by t^0.

Lemma 2: If $i \geqslant 0$ and if $x(s)$, $x(t)$ are any pair of edge labels in component i, then there is a unique element of $\mathbf{G}^*_{k,\nu}$ that interchanges $x(s)$ and $x(t)$.

Proof: This follows from Lemma 1, since the restriction of $\mathbf{G}^*_{k,\nu}$ to the edges in component i is just $\mathbf{G}_{k,\nu}$. □

A set S of error events is said to be *fixed* by $\mathbf{G}^*_{k,\nu}$ if for all $g \in \mathbf{G}^*_{k,\nu}$ and all $E \in S$ we have $g(E) \in S$.

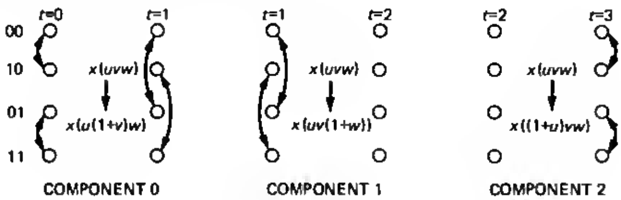

Fig. 4—Action of g_{t^0}, g_{t^1}, and g_{t^2}.

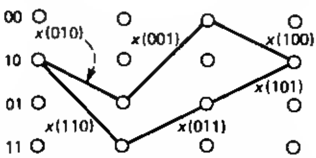

Fig. 5—The symmetry $(g_{t^0}, g_{t^1}, g_{t^2}, \cdots)$ applied to an error event.

Lemma 3: Let $i \geqslant 0$ and let S be a set of error events of the same length that is fixed by $\mathbf{G}^*_{k,\nu}$. If $m_i(x(a))$ is the total number of times the edge label $x(a)$ occurs in component i of the error events of S, then

$$m_i(x(a)) = \frac{2|S|}{2^{k+\nu}} \qquad \text{for all } \nu + k \text{ tuples } a.$$

Proof: Let s, t be binary $\nu + k$ tuples. By Lemma 2 there is an element of $\mathbf{G}^*_{k,\nu}$ interchanging error events involving $x(s)$ in component i with error events involving $x(t)$ in component i. Hence $m_i(x(s)) = m_i(x(t))$. Since the total number of edges in component i is $2|S|$, we have $m_i(x(a)) = 2|S|/2^{k+\nu}$ for all $\nu + k$ tuples a.

An *orbit* S of the group $\mathbf{G}^*_{k,\nu}$ is a set of error events satisfying

1. if $E \in S$ and $g \in \mathbf{G}^*_{k,\nu}$ then $g(E) \in S$, and
2. if $E_1, E_2 \in S$ then there exists $g \in \mathbf{G}^*_{k,\nu}$ such that $g(E_1) = E_2$.

Fig. 6 shows an orbit of $\mathbf{G}^*_{1,2}$. Observe that $m_i(x(a)) = 1$ for all i and for all a.

III. THE FIRST BOUND

In this section we derive the upper bound

$$\frac{d^2_{\min}}{P} \leqslant 4\left(1 + \frac{\nu}{k}\right).$$

This bound will be strengthened in later sections but it seems worth presenting the simpler argument here.

Observe that the average transmitted signal power is simply the average of the transmitted channel symbols, namely

$$P = \frac{1}{2^{\nu+k}} \sum_{i=0}^{2^{\nu+k}-1} x_i^2. \tag{18}$$

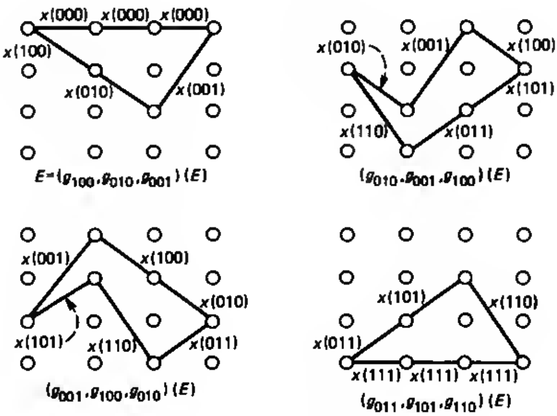

Fig. 6—An orbit of $\mathbf{G}^*_{1,2}$.

(Recall that the channel symbol $x(a_0 \cdots a_{\nu+k-1})$ is also denoted x_i where $i = a_0 + a_1 2^1 + \cdots + a_{\nu+k-1} 2^{\nu+k-1}$.) The Euclidean distance between the paths of the error event E sbown in Fig. 2 is

$$d^2(E) = (x_0 - x_1)^2 + (x_0 - x_3)^2 + (x_0 - x_6)^2,$$

which is a quadratic form in the variables x_i. In general we define

$$\underline{x}^T = (x_0, x_1, \cdots, x_{2^{\nu+k}-1}), \tag{19}$$

where the superscript T denotes matrix transpose. Then the Euclidean distance $d^2(E)$ between the paths of an error event E is given by

$$d^2(E) = \underline{x}^T A(E) \underline{x}, \tag{20}$$

where $A(E)$ is a symmetric, positive semi-definite matrix whicb we call the *distance matrix* of E. The distance matrix $A(E)$ has two properties that we wish to note:

Property I. The ith diagonal element of $A(E)$ counts the number of times the symhol x_i occurs in tbe error event.

Property II. The rows of $A(E)$ sum to zero.

By (18) and (20),

$$\frac{d^2_{\min}}{P} = \min_E \frac{\underline{x}^T A(E) \underline{x}}{P} = 2^{\nu+k} \min_E \frac{\underline{x}^T A(E) \underline{x}}{\underline{x}^T \underline{x}}, \tag{21}$$

where we minimize over all error events E.

Although we will make no use of the fact in this work, we note that in (21) only a finite number of error events need be considered, for no error event need be considered that bas a repeated pair of states. Thus, if the pair of states u and w occur at time i and also at a later time j, all components between i and j may he eliminated and the remainder of the error event after time j may he placed after time i. Since components cannot make a negative contrihution to $d^2(E)$ tbe new error event has distance no greater than the original one. By (21) the hest normalized minimum distance that can be acbieved for any choice of channel symbols is

$$2^{\nu+k} \max_{\underline{x}} \min_E \frac{\underline{x}^T A(E) \underline{x}}{\underline{x}^T \underline{x}}. \tag{22}$$

Consider an error event E with initial state (time $t = 0$) $a = (a_1, \cdots, a_{N-1})$ and final state $z = (z_1, \cdots, z_{N-1})$. If k tuples b_1, b_1^* are input at time 0, then at time 1 the two paths occupy states $(b_1, a_1, \cdots, a_{N-2})$ and $(b_1^*, a_1, \cdots, a_{N-2})$. There must be at least $N - 1$ further inputs before the paths can remerge. To remerge at z, tbe k tuples $z_{N-1}, z_{N-2}, \cdots, z_1$ must be input in that order to both paths. We denote this error event hy $E(a, z; b_1, b_1^*)$. Thus, the minimal length

of an error event is $N = (k + \nu)/k$. Fig. 2 shows the error event $E(00, 00; 0, 1)$ which has minimal length 3.

Given an arbitrary set S of error events, define

$$Q(S) = \frac{1}{|S|} \sum_{E \in S} A(E). \tag{23}$$

Let S^N be the set of all error events of length N. Note that S^N is fixed by the group $\mathbf{G}^*_{k,\nu}$.

Theorem 1: If k divides ν then the normalized minimum distance of any sliding window trellis code with 2^ν states and rate k bits/channel symbol satisfies

$$\frac{d^2_{\min}}{P} \leq 4\left(1 + \frac{\nu}{k}\right).$$

Proof: By (22),

$$\begin{aligned}\frac{d^2_{\min}}{P} &\leq 2^{\nu+k} \max_{\underline{x}} \min_{E} \frac{\underline{x}^T A(E) \underline{x}}{\underline{x}^T \underline{x}} \\ &\leq 2^{\nu+k} \max_{\underline{x}} \min_{E \in S^N} \frac{\underline{x}^T A(E) \underline{x}}{\underline{x}^T \underline{x}} \\ &\leq \frac{2^{\nu+k}}{|S^N|} \max_{\underline{x}} \frac{\underline{x}^T \left(\sum_{E \in S^N} A(E)\right) \underline{x}}{\underline{x}^T \underline{x}}.\end{aligned}$$

The last inequality simply states that the minimum is not more than the average. Setting $A_N = \sum_{E \in S^N} A(E)$, we have

$$\frac{2^{\nu+k}}{|S^N|} \max_{\underline{x}} \frac{\underline{x}^T A_N \underline{x}}{\underline{x}^T \underline{x}} = \frac{2^{\nu+k}}{|S^N|} \lambda_1(A_N),$$

where $\lambda_1(A_N)$ denotes the largest eigenvalue of A_N. By Property I, the ith diagonal entry of A_N counts the total number of times the edge x_i appears in some component of the error events of length N. By Lemma 3 all diagonal entries are equal to $2N|S^N|/2^{\nu+k}$. Property II implies that all row sums of A_N are zero. By the Gersgorin Circle Theorem[5]

$$\lambda_1(A_N) \leq 2(\text{diagonal entry}) = 2\left(\frac{2N|S^N|}{2^{\nu+k}}\right),$$

and so

$$\frac{d^2_{\min}}{P} \leq 4N = 4\left(1 + \frac{\nu}{k}\right). \quad \square$$

Remarks: In Section IV we derive a formula for $Q(S^N)$, and, by computing $\lambda_1(A_N)$, we prove

$$\frac{d^2_{\min}}{P} \leqslant \frac{2^{k+1}}{2^k - 1}\left(1 + \frac{\nu}{k}\right).$$

In Appendix B we prove that if $\nu = (N - 1)k + l$, where $0 \leqslant l < k$, then

$$\frac{d^2_{\min}}{P} \leqslant 4\left(1 + \left\lfloor\frac{\nu}{k}\right\rfloor\right),$$

where $\lfloor y \rfloor$ denotes the integer part of y.

IV. A FORMULA FOR $Q(S^N)$ AND A SHARPER BOUND

In this section we derive a formula for $Q(S^N)$, the matrix obtained by averaging the distance matrices of all error events of minimal length $N = (k + \nu)/k$. We require a matrix representation of the group $\mathbf{G}_{k,\nu}$.

If A is an $m \times n$ matrix and B is an $m_1 \times n_1$ matrix, then the tensor product $A \otimes B$ (also called the Kronecker product) is the $mm_1 \times nn_1$ matrix

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots\cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots\cdots & a_{2n}B \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_{m1}B & a_{m2}B & \cdots\cdots & a_{mn}B \end{bmatrix}.$$

Tensor products are discussed in Ref. 5, where they are called direct products. For appropriately sized matrices, A, B, C, and D, we have $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$. If λ is an eigenvalue of A with associated eigenvector v, and μ is an eigenvalue of B with associated eigenvector w, then $\lambda\mu$ is an eigenvalue of $A \otimes B$ with eigenvector $v \otimes w$.

We denote the $n \times n$ identity matrix by I_n and we abbreviate I_2 to I. Set

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \tag{24}$$

Consider the $2^{\nu+k} \times 2^{\nu+k}$ matrix

$$P_i = \underbrace{I \otimes \cdots \otimes I}_{j \text{ terms}} \otimes A \otimes \underbrace{I \otimes \cdots \otimes I}_{i \text{ terms}}$$

$$= I_{2^j} \otimes A \otimes I_{2^i}, \tag{25}$$

where $i + j + 1 = \nu + k$. This is the matrix

$$P_i = \begin{bmatrix} \begin{matrix} \mathbf{0} & I_{2^i} \\ I_{2^i} & \mathbf{0} \end{matrix} & & & \mathbf{0} \\ & \begin{matrix} \mathbf{0} & I_{2^i} \\ I_{2^i} & \mathbf{0} \end{matrix} & & \\ & & \ddots & \\ \mathbf{0} & & & \begin{matrix} \mathbf{0} & I_{2^i} \\ I_{2^i} & \mathbf{0} \end{matrix} \end{bmatrix}$$

with the indicated block repeated 2^j times along tbe main diagonal. Define u_i, $i = 0, 1, \cdots, 2^{\nu+k} - 1$, to be the binary $\nu + k$ tuple witb a 1 in position i and 0's elsewhere. Let

$$\underset{\sim}{x} = (x_0, \cdots, x_{2^{\nu+k}-1})^T = (x(0 \cdots 0), \cdots, x(1 \cdots 1))^T.$$

The permutation g_{u_i} maps $x(s)$ to $x(u_i + s)$ and so it interchanges edges with subscripts differing by 2^i. But this is precisely the effect of the transformation $\underset{\sim}{x} \rightarrow P_i\underset{\sim}{x}$. If t is an arbitrary $\nu + k$ tuple then the matrix describing the permutation g_t is obtained by multiplying the appropriate matrices P_i. For $t = (t_0, t_1, \cdots, t_{\nu+k-1})$ we define

$$M(t) = M_{\nu+k-1} \otimes \cdots \otimes M_1 \otimes M_0, \tag{26}$$

where

$$M_j = \begin{cases} I & \text{if} \quad t_j = 0 \\ A & \text{if} \quad t_j = 1. \end{cases} \tag{27}$$

Note that the subscript order in (26) is the reverse of the subscript order in the vector t. We have now proved the following lemma.

Lemma 4: If t is a $\nu + k$ tuple, then the permutation g_t: $x(s) \rightarrow x(s + t)$ is represented by $\underset{\sim}{x} \rightarrow M(t)\underset{\sim}{x}$.

As an example, the permutation g_{010} given in (12) is represented by the matrix $P = I \otimes A \otimes I$ given in (14). By Lemma 4 we may regard $\mathbf{G}_{k,\nu}$ as the following group of matrices:

$$\mathbf{G}_{k,\nu} = \{M_{\nu+k-1} \otimes \cdots \otimes M_1 \otimes M_0 \,|\, M_j = I \text{ or } A,$$
$$j = 0, \cdots, \nu + k - 1\}. \tag{28}$$

We shall prove that $Q(S^N)$ is a particular linear combination of

matrices $M(t)$ in $\mathbf{G}_{k,\nu}$. To calculate $\lambda_1(Q(S^N))$ we need to work with eigenvectors and eigenvalues of the matrices $M(t)$.

The matrices $M(t)$ are symmetric and they all commute; hence, they can be simultaneously diagonalized. Let H be the tensor product of $\nu + k$ copies of

$$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Observe that $H^{-1} = H^T$. Since $(1, 1)^T$ and $(1, -1)^T$ are eigenvectors of A, the columns of H are eigenvectors of $M(t)$ for all $\nu + k$ tuples t. Thus, $H^TM(t)H$ is diagonal for every matrix $M(t)$ in $\mathbf{G}_{k,\nu}$. If $p = (p_0, \cdots, p_{\nu+k-1})$ is a binary $\nu + k$ tuple, define

$$w(p) = w_{\nu+k-1} \otimes \cdots \otimes w_1 \otimes w_0,$$

where

$$w_j = \begin{cases} (1, 1)^T, & \text{if} \quad p_j = 0 \\ (1, -1)^T, & \text{if} \quad p_j = 1. \end{cases} \tag{29}$$

The vectors $w(p)$ are the columns of H. Note that $w(p)$ is formed by *reversing* the vector p. We have $A(1, 1)^T = (1, 1)^T$ and $A(1, -1)^T = -(1, -1)^T$. If $t = (t_0, \cdots, t_{\nu+k-1})$ then by (27) and (29)

$$\begin{aligned} M(t)w(p) &= \bigotimes_{j=0}^{\nu+k-1} M_j w_j = \left(\prod_{j=0}^{\nu+k-1} (-1)^{t_j p_j}\right) w(p) \\ &= (-1)^{p\cdot t} w(p), \end{aligned} \tag{30}$$

where $p \cdot t$ is the dot product of the vectors p and t.

Lemma 5: Suppose R is a diagonable matrix that commutes with every matrix $M(t)$ in $\mathbf{G}_{k,\nu}$. Then R is a linear combination of the matrices $M(t)$ in $\mathbf{G}_{k,\nu}$.

Proof: If s, t are different $\nu + k$ tuples, tben by Lemma 1, $g_s(x_0) \neq g_t(x_0)$. The permutation matrices $M(t)$ are therefore linearly independent because the 1's in row 0 are in different positions. Thus we have $2^{\nu+k}$ linearly independent diagonal matrices $H^{-1}M(t)H$. Since R commutes with every matrix $M(t)$, $H^{-1}RH$ commutes with every matrix $H^{-1}M(t)H$, and therefore $H^{-1}RH$ is diagonal. The matrices $H^{-1}M(t)H$ span the set of diagonal matrices so $H^{-1}RH$ is a linear combination of matrices $H^{-1}M(t)H$ and the lemma follows. □

Lemma 6: If S is a set of error events fixed by $\mathbf{G}^*_{k,\nu}$ then $(\sum_{E\in S} A(E))$ is a linear combination of the matrices $M(t)$ in $\mathbf{G}_{k,\nu}$.

Proof: The distance matrix $A(E)$ of an error event is the sum of contributions from each component:

$$A(E) = \sum_c A_c(E), \tag{31}$$

where we sum over the components c of E. The restriction of $\mathbf{G}^*_{k,\nu}$ to the edges in any component c is just the group $\mathbf{G}_{k,\nu}$. If edges $x(s)$, $x(\hat{s})$ appear in component c of error event E, then edges $g_t(x(s))$, $g_t(x(\hat{s}))$ appear in component c of error event $E' = g_t(E)$. We have

$$A_c(E') = M(t)^T A_c(E) M(t). \tag{32}$$

Since $M(t)$ is a permutation matrix and $M(t)^2 = I$, we have $M(t)^T = M(t)^{-1}$. Now g_t merely permutes the error events in S, so that by (32),

$$\sum_{E \in S} A_c(E) = \sum_{E \in S} A_c(g_t(E)) = \sum_{E \in S} M(t)^{-1} A_c(E) M(t)$$

$$= M(t)^{-1} \left(\sum_{E \in S} A_c(E) \right) M(t) \tag{33}$$

for all matrices $M(t)$ and for all components c. Summing (33) over all components c finishes the proof. $\square$

Example: If S is the orbit of error events shown in Fig. 6 then

$$\sum_{E \in S} A(E) = \left|\begin{array}{cc|cc|cc|cc} 3 & -1 & -1 & & -1 & & & \\ -1 & 3 & & -1 & & -1 & & \\ \hline -1 & & 3 & -1 & & & -1 & \\ & -1 & -1 & 3 & & & & -1 \\ \hline -1 & & & & 3 & -1 & -1 & \\ & -1 & & & -1 & 3 & & -1 \\ \hline & & -1 & & -1 & & 3 & -1 \\ & & & -1 & & -1 & -1 & 3 \end{array}\right|$$

$$= 3I \otimes I \otimes I - (I \otimes I \otimes A + I \otimes A \otimes I + A \otimes I \otimes I). \tag{34}$$

Consider S^N, the set of all error events $E(a, z; b_1, b_1^*)$ of minimal length $N = (k + \nu)/k$. Recall that $a = (a_1, \cdots, a_{N-1})$ is the initial state, $z = (z_1, \cdots, z_{N-1})$ is the final state, and b_1, b_1^* are the first pair of inputs. We have $|S^N| = \binom{2^k}{2} 2^\nu \cdot 2^\nu$.

Lemma 7:
(1) *Let* $t = (t_0, \cdots, t_{N-1})$ *and let* $t' = (t_1, \cdots, t_{N-1})$ *where* t_i, $i = 0, 1, \cdots, N-1$, *is a binary k tuple. If* $g_t^* = (g_{t^0}, g_{t^1}, \cdots, g_{t^{N-1}}) \in \mathbf{G}^*_{k,\nu}$, *then*

$$g_t^*(E(a, z; b_1, b_1^*)) = E(a + t', z + t', b_1 + t_0, b_1^* + t_0). \tag{35}$$

(2) *The group* $\mathbf{G}^*_{k,\nu}$ *partitions the set* S^N *of error events of length* N *into* $2^\nu(2^k - 1)$ *orbits each of size* $2^{\nu+k-1}$.

Proof: Part (1) follows from the definition of g_{t^i} given in (17). To verify part (2) we note that $E(a, z; b_1, b_1^*)$ is fixed only by the symmetry g_b^*, where $b = (b_1 + b_1^*, 0, 0, \cdots, 0)$. Hence, each orbit consists of

$2^{\nu+k-1}$ distinct error events. Since the total number of error events in S^N is $(2^\nu \cdot 2^\nu(2^k)(2^k-1))/2$, we see that there are $2^\nu(2^k-1)$ orbits. □

The orbit containing the error event $E(a, z; b_1, b_1^*)$ is determined by $a + z$ and $b_1 + b_1^*$. Setting $f = b_1 + b_1^*$, we denote this orbit by $S(a+z; f)$. This orbit contains $E(a, z; 0, f)$; note that $f \neq 0$ because $b_1 \neq b_1^*$. Recall that if f is a k tuple, then the $\nu + k$ tuple $(f0 \cdots 0)^i$ equals $(y_0, y_1, \cdots, y_{N-1})$ where $y_i = f$ and $y_j = 0$ for $j \neq i$.

Lemma 8: Let S^N be the set of all error events of length N and let $S(a+z; f)$ be the orbit of $\mathbf{G}_{k,\nu}^*$ containing the error event $E(a, z; 0, f)$. Then

$$(1)\ 2^{\nu+k}Q(S(a+z; f)) = 2NI_{2^{\nu+k}} - 2\sum_{i=0}^{N-1} M((f0 \cdots 0)^i) \tag{36}$$

$$(2)\ 2^{\nu+k}(2^k-1)Q(S^N) = 2(2^k-1)NI_{2^{\nu+k}} - 2\sum_{f\neq 0}\sum_{i=0}^{N-1} M((f0 \cdots 0)^i). \tag{37}$$

Proof: We calculate the contribution to $Q(S(a+z; f))$ made by pairs of edges in component 0. Since the restriction of $\mathbf{G}_{k,\nu}^*$ to the edges in any component is just the group $\mathbf{G}_{k,\nu}$, this distance contribution is

$$\frac{1}{2^{\nu+k}}\sum_t [g_t(x(0a_1 \cdots a_{N-1})) - g_t(x(fa_1 \cdots a_{N-1}))]^2$$

$$= \frac{1}{2^{\nu+k}}\sum_t [x(t + (0a_1 \cdots a_{N-1})) - x(t + (fa_1 \cdots a_{N-1}))]^2$$

$$= \frac{1}{2^{\nu+k}}\left[2\sum_t x(t)^2 - 2\sum_t x(t)x(t + (f0 \cdots 0))\right]$$

$$= \frac{1}{2^{\nu+k}}\,\underline{x}^T[2I_{2^{\nu+k}} - 2M(f0 \cdots 0)]\underline{x}.$$

In general, the distance contribution made by edges in component i is

$$\frac{1}{2^{\nu+k}}\sum_t [g_t(x(z_{N-i} \cdots z_{N-1}0a_1 \cdots a_{N-i-1})) - g_t(x(z_{N-i} \cdots z_{N-1}fa_1 \cdots a_{N-i-1}))]^2$$

$$= \frac{1}{2^{\nu+k}}\sum_t [x(t + (z_{N-i} \cdots z_{N-1}0a_1 \cdots a_{N-i-1})) - x(t + (z_{N-i} \cdots z_{N-1}fa_1 \cdots a_{N-i-1}))]^2$$

$$= \frac{1}{2^{\nu+k}}\,\underline{x}^T[2I_{2^{\nu+k}} - 2M((f0 \cdots 0)^i)]\underline{x}. \tag{38}$$

Summing (38) over all components i, we obtain (36). Since (36) is independent of $a + z$, we obtain tbe formula for $Q(S^N)$ by summing (36) over all nonzero k tuples f. □

Remark: When $k = 1$, there is only one cboice for f, namely $f = 1$, and so every form $Q(S(a + z; f))$ is equal to $Q(S^N)$. For $k = 1$, $\nu = 2$, we have

$$Q(S(00; 1)) = Q(S(10; 1)) = Q(S(01; 1)) = Q(S(11; 1)) = Q(S^3)$$
$$= 1/4[3I_8 - (I \otimes I \otimes A + I \otimes A \otimes I + A \otimes I \otimes I)]$$

[see the matrix given as (34)]. However, for $k > 1$, the form $Q(S(a + z; f))$ will change with f. Thus, for $k = 2$, $\nu = 4$, we have

$$Q(S(a + z; 11)) = 1/32[3I_{64} - (I_4 \otimes I_4 \otimes (A \otimes A)$$
$$+ I_4 \otimes (A \otimes A) \otimes I_4 + (A \otimes A) \otimes I_4 \otimes I_4)],$$

while

$$Q(S(a + z; 10)) = 1/32[3I_{64} - (I_4 \otimes I_4 \otimes (I \otimes A)$$
$$+ I_4 \otimes (I \otimes A) \otimes I_4 + (I \otimes A) \otimes I_4 \otimes I_4)].$$

Theorem 2: If k divides ν, then the normalized minimum distance of any sliding window trellis code with 2^ν states and rate k bits/channel symbol satisfies

$$\frac{d^2_{\min}}{P} \leqslant \frac{2^{k+1}}{2^k - 1}\left(1 + \frac{\nu}{k}\right).$$

Proof: From the proof of Theorem 1, we have

$$\frac{d^2_{\min}}{P} \leqslant 2^{\nu+k}\lambda_1[Q(S^N)]$$
$$= \frac{1}{2^k - 1}\lambda_1(Q_N), \tag{39}$$

where $Q_N = (2^k - 1)2^{\nu+k}Q(S^N)$. Let $c = (c_0, \cdots, c_{N-1})$ be a binary $\nu + k$ tuple and let γ be the number of nonzero k tuples c_i. Then by (30), the eigenvalue of Q_N associated with $w(c)$ is

$$2(2^k - 1)N - 2\sum_{f\neq 0}\sum_{i=0}^{N-1}(-1)^{c_i \cdot f}$$
$$= 2(2^k - 1)N - 2\sum_{i=0}^{N-1}\sum_{f\neq 0}(-1)^{c_i \cdot f}, \tag{40}$$

where we sum over all nonzero k tuples f. Since

$$\sum_{f\neq 0} (-1)^{c_i \cdot f} = \begin{cases} 2^k - 1, & \text{if} \quad c_i = 0 \\ -1, & \text{if} \quad c_i \neq 0, \end{cases}$$

eq. (40) becomes

$$2(2^k - 1)N - 2[(2^k - 1)(N - \gamma) - \gamma] = 2^{k+1}\gamma. \tag{41}$$

The largest eigenvalue of Q_N is obtained when $\gamma = N = (1 + (\nu/k))$. The theorem now follows from (39). □

Remarks: Observe tbat the largest eigenvalue of $Q(S^N)$ is associated with $w(\mathrm{c})$, where $c_0, c_1, \cdots, c_{N-1}$ are all nonzero. For example, with $k = 1$, the largest eigenvalue, $4(1 + (\nu/k))$ bas multiplicity one and is associated with tbe eigenvector $(1, -1)^T \otimes (1, -1)^T \otimes \cdots \otimes (1, -1)^T$. When $k > 1$, there will be several linearly independent eigenvectors associated with $\lambda_1(Q(S^N))$ because tbere are several cboices for c with all $c_i \neq 0$. Also, note that Theorem 2 gives the same bound as Theorem 1 when $k = 1$. For $k \geqslant 2$, the bound of Theorem 2 is an improvement.

In Appendix B we prove that if $\nu = (N - 1)k + l$ where $0 \leqslant l < k$, then

$$\frac{d_{\min}^2}{P} \leqslant \frac{2^{k-l+1}}{2^{k-l} - 1}\left(1 + \left\lfloor \frac{\nu}{k} \right\rfloor\right),$$

where $\lfloor y \rfloor$ denotes the integer part of y.

V. A FINAL BOUND OBTAINED FROM A WEIGHTED AVERAGE

Let S^{N+1} be the set of all error events of length $N + 1 = 2 + (\nu/k)$. Let $Q(S^{N+1})$ be the matrix obtained by averaging the distance matrices of all error events of length $N + 1$. In this section we derive a formula for $Q(S^{N+1})$ and we prove

$$\frac{d_{\min}^2}{P} \leqslant \frac{2^{2k+1}}{2^{2k} - 1}\left(2 + \frac{\nu}{k}\right)$$

using a weighted average of $Q(S^N)$ and $Q(S^{N+1})$.

An error event E of length $N + 1$ is determined by the initial state $a = (a_1, \cdots, a_{N-1})$, the final state $z = (z_1, \cdots, z_{N-1})$, the inputs b_1, b_1^* at time 0, and the inputs b_2, b_2^* at time 1. Since the two paths diverge at time 0, we must have $b_1 \neq b_1^*$. To remerge at z the last $N - 1$ inputs must be the k tuples $z_{N-1}, z_{N-2}, \cdots, z_1$ in that order. After N inputs the two paths occupy states $z_2 \cdots z_{N-1}b_2$ and $z_2 \cdots z_{N-1}b_2^*$. At this stage the two paths must be disjoint so $b_2 \neq b_2^*$. We denote this error event E by $E(a, z; b_1, b_1^*; b_2, b_2^*)$ [equivalently $E(a, z; b_1^*, b_1; b_2^*, b_2)$].

The group $\mathrm{G}_{k,\nu}^*$ maps error events of length $N + 1$ to error events of

length $N + 1$. To he specific, let $t = (t_0, \cdots, t_{N-1})$ he a $\nu + k$ tuple and set $t' = (t_1, \cdots, t_{N-1})$, $t'' = (t_0, \cdots, t_{N-2})$. If $g_t^* = (g_t, g_{t^1}, \cdots, g_{t^{N-1}}, g_t)$ then it follows from the definition of g_{t^i} given in (7) that

$$g_t^*(E(a, z; b_1, b_1^*; b_2, b_2^*))$$
$$= E(a + t', z + t''; b_1 + t_0, b_1^* + t_0; b_2 + t_{N-1}, b_2^* + t_{N-1}). \quad (42)$$

This group action does not preserve $a + z$ hut it does preserve $b_1 + b_1^*$ and $b_2 + b_2^*$. Set $g = b_1 + b_1^*$ and $f = b_2 + b_2^*$. We denote the orhit of $\mathbf{G}_{k,\nu}^*$ containing the error event $E(a', z'; 0, g; 0, f)$ hy $S(a', z'; g, f)$ (in the discussion ahove, $a' = (a_1, \cdots, a_{N-2}, a_{N-1} + b_2)$ and $z' = (z_1 + b_1, z_2, \cdots, z_{N-1})$). Note that $f, g \neq 0$.

If f, g are k tuples, then the $\nu + k$ tuple $(fg0 \cdots 0)^0 = (fg0 \cdots 0)$ and $(fg0 \cdots 0)^i$ is ohtained from $(fg0 \cdots 0)^{i-1}$ hy cycling the hlocks of k hits to the right and moving the last hlock to the front. Thus $(fg0 \cdots 0)^{N-2} = (0 \cdots 0fg)$. Define matrices $M_i(fg0 \cdots 0)$, $i = 0, \cdots, N$ in $\mathbf{G}_{k,\nu}$ as follows:

$$\begin{aligned} M_0(fg0 \cdots 0) &= M(g0 \cdots 0) \\ M_i(fg0 \cdots 0) &= M((fg0 \cdots 0)^{i-1}) \qquad i = 1, \cdots, N-1, \\ M_N(fg0 \cdots 0) &= M(0 \cdots 0f). \end{aligned} \quad (43)$$

Example: For $k = 1$, $\nu = 2$, the orhit $S(00, 00; 1, 1)$ is shown in Fig. 7. The quadratic form $Q(S(00, 00; 1, 1))$ is given hy

$$Q(S(00, 00; 1, 1)) = \frac{1}{8} \left|\begin{array}{cc|cc|cc|cc} 8 & -2 & & -2 & -2 & & -2 & \\ -2 & 8 & -2 & & & -2 & & -2 \\ \hline & -2 & 8 & -2 & -2 & & -2 & \\ -2 & & -2 & 8 & & -2 & & -2 \\ \hline -2 & & -2 & & 8 & -2 & & -2 \\ & -2 & & -2 & -2 & 8 & -2 & \\ \hline -2 & & -2 & & & -2 & 8 & -2 \\ & -2 & & -2 & -2 & & -2 & 8 \end{array}\right|$$

$$= \frac{1}{8}(8I_8 - 2(I \otimes I \otimes A + I \otimes A \otimes A + A \otimes A \otimes I$$
$$+ A \otimes I \otimes I))$$
$$= \frac{1}{8}\left(8I_8 - 2\sum_{i=0}^{3} M_i(110)\right). \quad (44)$$

Lemma 9: Let S^{N+1} be the set of all error events of length $N + 1$ and let $S(a, z; g, f)$ be the orbit of $\mathbf{G}_{k,\nu}^*$ containing the error event $E(a, z; 0, g; 0, f)$. Then,

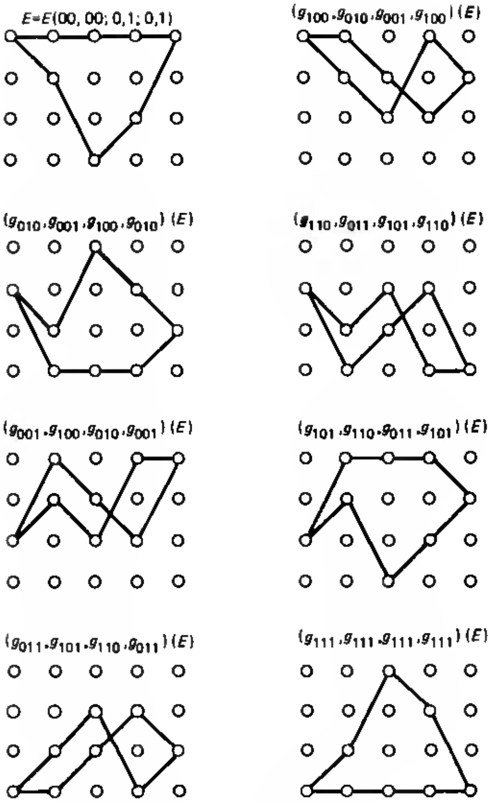

Fig. 7—The orbit $S(00, 00; 1, 1)$.

$$(1)\ 2^{\nu+k}Q(S(a, z; g, f)) = 2(N + 1)I_{2^{\nu+k}} - 2 \sum_{i=0}^{N} M_i(fg0 \cdots 0). \tag{45}$$

$$(2)\ 2^{\nu+k}(2^k - 1)^2 Q(S^{N+1}) = 2(2^k - 1)^2(N + 1)I_{2^{\nu+k}} - 2 \sum_{f,g \neq 0} \sum_{i=0}^{N} M_i(fg0 \cdots 0). \tag{46}$$

Proof: We calculate the contribution to $Q(S(a, z; g, f))$ made by pairs of edges in component 0. This distance contribution is

$$\frac{1}{2^{\nu+k}} \sum_t [x(t + (0a_1 \cdots a_{N-1})) - x(t + (ga_1 \cdots a_{N-1}))]^2$$

$$= \frac{1}{2^{\nu+k}} \underline{x}^T[2I_{2^{\nu+k}} - 2M(g0 \cdots 0)]\underline{x}$$

as found in the proof of Lemma 8. Similarly, the contribution made

by pairs of edges in component N (the last component of the error events) is

$$\frac{1}{2^{\nu+k}} \sum_t [x(t + (z_1 \cdots z_{N-1}0)) - x(t + (z_1 \cdots z_{N-1}f))]^2$$
$$= \frac{1}{2^{\nu+k}} \underset{\sim}{x}^T[2I_{2^{\nu+k}} - 2M(0 \cdots 0f)]\underset{\sim}{x}.$$

For $i = 1, \cdots, N - 1$, the contribution made by pairs of edges in component i is

$$\frac{1}{2^{\nu+k}} \sum_t [x(t + (z_{N-i+1} \cdots z_{N-1}00a_1 \cdots a_{N-i-1}))$$
$$- x(t + (z_{N-i+1} \cdots z_{N-1}fga_1 \cdots a_{N-i-1}))]^2$$
$$= \frac{1}{2^{\nu+k}}\left[2 \sum_t x(t)^2 - 2 \sum_t x(t)x(t + (fg0 \cdots 0)^{i-1})\right]$$
$$= \frac{1}{2^{\nu+k}} \underset{\sim}{x}^T[2I_{2^{\nu+k}} - 2M((fg0 \cdots 0)^{i-1})]\underset{\sim}{x}.$$

The sum of the contributions from all $N + 1$ components is

$$Q(S(a, z; g, f)) = \frac{1}{2^{\nu+k}} \underset{\sim}{x}^T\left[2(N + 1)I_{2^{\nu+k}} - 2 \sum_{i=0}^{N} M_i(fg0 \cdots 0)\right]\underset{\sim}{x}.$$

This proves part (1). Observe that (45) is independent of a and z. We obtain $Q(S^{N+1})$ by summing (45) over all pairs g, f of nonzero k tuples. Since there are $(2^k - 1)^2$ such pairs,

$$(2^k - 1)^2 2^{\nu+k} Q(S^{N+1})$$
$$= 2(2^k - 1)^2(N + 1)I_{2^{\nu+k}} - 2 \sum_{f,g\neq 0} \sum_{i=0}^{N} M_i(fg0 \cdots 0)$$

as required. □

Remarks: When $k = 1$, we must have $f = g = 1$ and so every form $Q(S(a, z; g, f))$ is equal to $Q(S^{N+1})$. In this case, $N = 1 + \nu$ and

$$Q(S^{N+1}) = \frac{1}{2^{\nu+1}}\left[2(2 + \nu)I_{2^{\nu+1}} - 2 \sum_{i=0}^{1+\nu} M_i(110 \cdots 0)\right]$$

[see the matrix given as (44)]. For $k > 1$, there are several choices for f and g. Thus, for $k = 2$, $\nu = 4$, we have, with $g = (1, 1)$ and $f = (0, 1)$,

$$Q(S(a, z; 11, 01)) = 1/64[8I_{64} - 2(I_4 \otimes I_4 \otimes (A \otimes A)$$
$$+ I_4 \otimes (A \otimes A) \otimes (A \otimes I)$$
$$+ (A \otimes A) \otimes (A \otimes I) \otimes I_4$$
$$+ (A \otimes I) \otimes I_4 \otimes I_4)],$$

while with $g = (01)$ and $f = (10)$ we get

$$\begin{aligned} Q(S(a, z; 01; 10)) = 1/64[8I_{64} - 2(&I_4 \otimes I_4 \otimes (A \otimes I) \\ &+ I_4 \otimes (A \otimes I) \otimes (I \otimes A) \\ &+ (A \otimes I) \otimes (I \otimes A) \otimes I_4 \\ &+ (I \otimes A) \otimes I_4 \otimes I_4)]. \end{aligned}$$

Theorem 3: If k divides ν then the normalized minimum distance of any convolutionally derived trellis code with 2^ν states and rate k bits/channel symbol satisfies

$$\frac{d_{\min}^2}{P} \leqslant \frac{2^{2k+1}}{2^{2k} - 1}\left(2 + \frac{\nu}{k}\right).$$

Proof: If $\bar{Q}$ is any weighted average of $Q(S^N)$ and $Q(S^{N+1})$, then by (8) we have

$$\frac{d_{\min}^2}{P} \leqslant 2^{\nu+k}\lambda_1(\bar{Q}). \tag{47}$$

Let $\delta = 1/(2^{2k} - 1)$. Then $2(2^k - 1)\delta + (2^k - 1)^2\delta = 1$. Define $\bar{Q}$ to be the following weighted average of $Q(S^N)$ and $Q(S^{N+1})$:

$$\bar{Q} = 2(2^k - 1)\delta Q(S^N) + (2^k - 1)^2\delta Q(S^{N+1}).$$

Set

$$Q_N = 2^{\nu+k}(2^k - 1)Q(S^N)$$

and

$$Q_{N+1} = 2^{\nu+k}(2^k - 1)^2 Q(S^{N+1}).$$

Then by (47)

$$\frac{d_{\min}^2}{P} \leqslant \delta\lambda_1(2Q_N + Q_{N+1}). \tag{48}$$

The eigenvectors, $w(c)$, of Q_N and Q_{N+1} are in 1-1 correspondence with binary vectors $c = (c_1, \cdots, c_N)$, where c_i, $i = 1, \cdots, N$ are k tuples. By (41)

$$Q_N w(c) = 2^{k+1}\gamma(c)w(c), \tag{49}$$

where $\gamma(c)$ is the number of nonzero k tuples c_i. Introduce k tuples $c_0 = c_{N+1} = 0$ and define

$$\alpha(c) = |\{i \mid c_i = 0,\ c_{i+1} \neq 0 \text{ or } c_i \neq 0,\ c_{i+1} = 0\}|$$

and

$$\beta(c) = |\{i \mid c_i \neq 0 \text{ and } c_{i+1} \neq 0\}|. \tag{50}$$

There are $(N + 1) - \alpha(c) - \beta(c)$ indices i, $0 \leq i \leq N$, for which $c_i = c_{i+1} = 0$. By (30) the eigenvalue of Q_{N+1} associated with $w(c)$ is

$$2(2^k - 1)^2(N + 1) - 2 \sum_{f,g\neq 0} \sum_{i=0}^{N} (-1)^{c_i \cdot f + c_{i+1} \cdot g}$$

$$= 2(2^k - 1)^2(N+1) - 2 \sum_{i=0}^{N} \left(\sum_{f\neq 0} (-1)^{c_i \cdot f} \right) \left(\sum_{g \neq 0} (-1)^{c_{i+1} \cdot g} \right). \tag{51}$$

Recall that the sum $\sum_{f\neq 0} (-1)^{c_i \cdot f}$ is $(2^k - 1)$ when $c_i = 0$, hut equal to -1 whenever $c_i \neq 0$. Hence (50) is equal to

$$\begin{aligned} &2(2^k - 1)^2(N + 1) - 2[((N + 1) - \alpha(c) - \beta(c))(2^k - 1)^2 \\ &\qquad\qquad - \alpha(c)(2^k - 1) + \beta(c)] \\ &= 2(2^k - 1)^2(\alpha(c) + \beta(c)) + 2(2^k - 1)\alpha(c) - 2\beta(c) \\ &= 2^{k+1}(2^k(\alpha(c) + \beta(c)) - 2(\alpha(c) + \beta(c)) + \alpha(c)) \\ &= 2^{k+1}[2^k(\alpha(c) + \beta(c)) - (\alpha(c) + 2\beta(c))]. \end{aligned} \tag{52}$$

Now, $\gamma(c)$ is the numher of nonzero c_i's. Since each nonzero c_i appears in two pairs, (c_{i-1}, c_i) and (c_i, c_{i+1}), we have

$$\alpha(c) + 2\beta(c) = 2\gamma(c). \tag{53}$$

Suhstitution in (52) shows that the eigenvalue in (51), of Q_{N+1} associated with $w(c)$ is

$$2^{k+1}[2^k(2\gamma(c) - \beta(c)) - 2\gamma(c)]. \tag{54}$$

By (49) and (54) we have

$$\begin{aligned} &(2Q_N + Q_{N+1})w(c) \\ &\quad = 2^{k+1}(2\gamma(c) + 2^k(2\gamma(c) - \beta(c)) - 2\gamma(c))w(c) \\ &\quad = 2^{2k+1}(2\gamma(c) - \beta(c))w(c). \end{aligned} \tag{55}$$

There are $N - \gamma(c)$ indices i, $1 \leq i \leq N$, for which $c_i = 0$. Since every c_j, $1 \leq j \leq N$, appears in the two pairs (c_{j-1}, c_j) and (c_j, c_{j+1}), there are at most $2 + 2(N - \gamma(c))$ indices i, $0 \leq i \leq N$, for which $c_i = 0$ or $c_{i+1} = 0$. Hence

$$\beta(c) \geq (N + 1) - 2 - 2(N - \gamma(c)) = 2\gamma(c) - N - 1$$

and

$$2(\gamma(c)) - \beta(c) \leq N + 1. \tag{56}$$

Now (48), (55), and (56) imply

$$\frac{d_{\min}^2}{P} \leqslant \frac{2^{2k+1}}{2^{2k}-1}(N+1) = \frac{2^{2k+1}}{2^{2k}-1}\left(2+\frac{\nu}{k}\right). \quad \square$$

Remarks: Equality can hold in (56). If N is odd, set $c_0 = c_2 = \cdots = c_{N+1} = 0$ and $c_1, c_3, \cdots, c_N \neq 0$. Then $\gamma(c) = (N+1)/2$ and $\beta(c) = 0$. (Observe that for $k = 1$, $\nu = 2$, the largest eigenvalue of the form $2Q_N + Q_{N+1}$ is associated with eigenvector $(1, -1)^T \otimes (1, 1)^T \otimes (1, -1)^T$.) If N is even, set $c_0 = c_3 = c_5 = c_7 = \cdots = c_{N+1} = 0$ and $c_1, c_2, c_4, c_6, \cdots, c_N \neq 0$ to get $\gamma(c) = (N+2)/2$ and $\beta(c) = 1$. Setting

$$\frac{2^{k+1}}{2^k-1}\left(1+\frac{\nu}{k}\right) = \frac{2^{2k+1}}{2^{2k}-1}\left(2+\frac{\nu}{k}\right)$$

yields $\nu = k(2^k - 1)$. If $\nu < k(2^k - 1)$, then Theorem 2 gives the stronger bound; if $\nu > k(2^k - 1)$ then Theorem 3 gives the stronger bound. In particular, for $k = 1$, Theorem 3 gives a stronger bound for any $\nu > 1$.

The bound given by Theorem 3 is obtained from the largest eigenvalue of a particular weighted average of $Q(S^N)$ and $Q(S^{N+1})$. In Appendix A we use the duality theorem of linear programming to prove that no other weighted average of $Q(S^N)$ and $Q(S^{N+1})$ gives a stronger bound.

In Appendix B we prove that if $\nu = (N-1)k + l$, where $0 \leqslant l < k$, then

$$\frac{d_{\min}^2}{P} \leqslant \frac{2^{2(k-l)+1}}{2^{2(k-l)}-1}\left(2+\left\lfloor\frac{\nu}{k}\right\rfloor\right),$$

where $\lfloor y \rfloor$ denotes the integer part of y.

VI. CONCLUSIONS

Three upper bounds on the normalized minimum distance, $(d_{\min}^2/P)$, have been given for trellis codes. The bound

$$\frac{d_{\min}^2}{P} \leqslant 4\left(1+\frac{\nu}{k}\right)$$

given in Theorem 1 is typical. This certainly provides nontrivial information. For example, is it possible to gain 10 dB in minimum distance using $2^6 = 64$ states at rate 1 bit/symbol? The answer is no. Theorem 1 bounds the gain at 8.4 dB; Theorem 3 bounds the gain at 7.3 dB. Nevertheless, there still remain the questions of how tight these bounds are and if they exhibit the "right" dependence on the

Table I—Possible gains at rate 1 bit/symbol

ν	Lower Bound (Ungerboeck)	Upper Bounds: Theorems 1 and 2	Upper Bounds: Theorem 3
2	2.5 db	4.7 db	4.3 db
3	3	6.0	5.2
4	3.4	7.0	6.0
5	4.2	7.8	6.7
6	4.5	8.4	7.3
7	5.1	9.0	7.8
8	5.3	9.5	8.2
9	5.6	10.0	8.6
10	5.8	10.4	9.0
11	6	10.7	9.4

parameters ν and k. For example, consider the normalized minimum distance for block codes of length n, having 2^{nk} code words (k bits/symbol). In that case, known upper bounds bebave, for large n, like $d^2/P \lesssim 2n/4^k$. Tbus the linear dependence on ν, a quantity analogous to block length, appears correct. However, tbe true dependence on k may be different from our bound. Table I gives upper and lower bounds on the gain (in dB) that is possible at rate 1 bit/cbannel symbol. Tbe lower bounds arise from codes constructed by Ungerboeck.[4]

Also minimum distance is by no means the complete story with regard to error rate. The heuristics leading to tbe claim that terms involving $d_{\min}$ would dominate an upper bound on tbe error rate make the assumption that the infinite series determining tbe upper bound converges. Even if a code with a good $d_{\min}$ were found, an upper bound on error rate sbould still be computed for that particular code. As an example of a catastrophe that may occur, consider the assignment of edge labels $\underline{x}^T = (1, -1, -1, 1, -1, 1, 1, -1)$ to the trellis of Fig. 1. One observes that a pair of edges leaving a node always contributes $(1 - (-1))^2 = 4$ to tbe distance and similarly for a pair of edges merging into a node. One immediately concludes tbat no error event bas distance less than 8 for this edge assignment. Since $P = 1$, this is a 3 dB gain over the uncoded ± 1 situation. How could this happen witb only ± 1 symbols? One answer is that we forgot to include unmerged events, events which go on forever. We had implicitly assigned infinity to their distance, but now some bave distance 4. However, this could be rectified by perturbing the ± 1 edge labels by small amounts. A more serious trouble with this code is tbat an infinite number of error events have (essentially) the minimum distance and so a coefficient that we did not explicitly consider turns out to be infinite for this particular code.

REFERENCES

1. A. D. Wyner, "On Coding and Information Theory," SIAM Review, *11*, No. 3 (July 1969), pp. 317–46.
2. F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting-Codes*, Amsterdam: North-Holland and New York: Elsevier/North-Holland, 1977.
3. G. D. Forney, "The Viterbi Algorithm," Proc. IEEE, *61*, No. 3 (March 1973), pp. 268–78.
4. G. Ungerboeck, "Channel Coding with Multilevel/Phase Signals," IEEE Trans. Inform. Theory, *IT-28*, No. 1 (January, 1982), pp. 55–67.
5. Peter Lancaster, *Theory of Matrices*, New York: Academic Press, 1969.

APPENDIX A

The upper bound of Theorem 3 is obtained from the largest eigenvalue of a particular weighted average of the quadratic forms $Q(S^N)$ and $Q(S^{N+1})$. In this appendix we prove tbat no other weigbted average gives a stronger bound. We shall assume throughout that $\nu \geqslant k(2^k - 1)$ since tbe bound given in Theorem 3 improves upon tbat given in Theorem 2 only for ν in this range.

If $r_1, r_2 \geqslant 0$ and $r_1 + r_2 = 1$, then

$$\frac{d^2_{\min}}{P} \leqslant 2^{\nu+k}\lambda_1(r_1 Q(S^N) + r_2 Q(S^{N+1})).$$

Recall from (29) that the eigenvectors $w(c)$ of $Q(S^N)$ and $Q(S^{N+1})$ are 1:1 correspondence with binary $\nu + k$ tuples c. Let $c = (c_1, \cdots, c_N)$, where c_i, $i = 1, \cdots, N$ is a binary k tuple and let $c_0 = c_{N+1} = 0$. Recall that

$$\alpha(c) = |\{i \,|\, c_i = 0,\ c_{i+1} \neq 0 \text{ or } c_i \neq 0,\ c_{i+1} = 0\}|,$$

$$\beta(c) = |\{i \,|\, c_i \neq 0 \text{ and } c_{i+1} \neq 0\},$$

and

$$\gamma(c) = |\{i \,|\, c_i \neq 0\}|.$$

Define $\phi_N(c)$ and $\phi_{N+1}(c)$ by $2^{\nu+k}(2^k - 1)Q(S^N)w(c) = \phi_N(c)w(c)$ and $2^{\nu+k}(2^k - 1)^2 Q(S^{N+1})w(c) = \phi_{N+1}(c)w(c)$. Then by (49) and (54)

$$\phi_N(c) = 2^{k+1}\gamma(c) \tag{57}$$

and

$$\phi_{N+1}(c) = 2^{k+1}[2^k(2\gamma(c) - \beta(c)) - 2\gamma(c)]. \tag{58}$$

To find the optimal weighted average we have to solve the following linear programming problem.

Choose real variables $r_1, r_2, r \geqslant 0$ so as to minimize r subject to the inequalities

$$-(r_1 + r_2) \leqslant -1 \tag{59}$$

and

$$r_1 \frac{\phi_N(c)}{2^k - 1} + r_2 \frac{\phi_{N+1}(c)}{(2^k - 1)^2} - r \leqslant 0, \qquad \text{for all} \quad \nu + k \text{ tuples } c.$$

In Theorem 3 we proved that a feasible solution to (43) is

$$r_1 = \frac{2}{2^k + 1}, \qquad r_2 = \frac{2^k - 1}{2^k + 1}, \qquad r = \frac{2^{2k+1}}{2^{2k} - 1}\left(2 + \frac{\nu}{k}\right). \tag{60}$$

The linear program (59) is the dual of the primal linear program given below.

Choose real variables a_c, $a \geqslant 0$, where the index c runs through all binary $\nu + k$ tuples, so as to maximize a subject to the inequalities

$$\frac{1}{2^k - 1}\left(\sum_c \phi_N(c)a_c\right) - a \geqslant 0$$

$$\frac{1}{(2^k - 1)^2}\left(\sum_c \phi_{N+1}(c)a_c\right) - a \geqslant 0$$

$$-\left(\sum_c a_c\right) \geqslant -1. \tag{61}$$

If we can find a feasible solution to (61) with

$$a = \frac{2^{2k+1}}{2^{2k} - 1}\left(2 + \frac{\nu}{k}\right),$$

then by the duality theorem of linear programming,[2] (60) is an optimal solution to (59). We consider two cases.

Case 1. N odd

Pick $f = (f_1, \cdots, f_N)$, where f_i, $i = 1, \cdots, N$ is a binary k tuple and every f_i is nonzero. Pick $g = (g_1, \cdots, g_N)$, where g_i, $i = 1, \cdots, N$ is a binary k tuple and $g_i \neq 0$ if and only if i is odd. Then $\gamma(f) = N$, $\beta(f) = N - 1$ and $\gamma(g) = (N + 1)/2$, $\beta(g) = 0$. By (57) and (58), $\phi_N(f)$, $\phi_N(g)$, $\phi_{N+1}(f)$, and $\phi_{N+1}(g)$ are as follows:

	f	g
ϕ_N	$2^{k+1}N$	$2^k(N + 1)$
ϕ_{N+1}	$2^{k+1}[2^k(N + 1) - 2N]$	$2^{k+1}[(2^k - 1)(N + 1)]$

Set

$$a_c = \begin{cases} \left(\dfrac{2^k - 1}{2^k + 1}\right)\left(\dfrac{N + 1}{N - 1}\right), & \text{if} \quad c = f \\ 1 - a_f = \dfrac{2(N - 2^k)}{(2^k + 1)(N - 1)}, & \text{if} \quad c = g \\ 0, & \text{otherwise,} \end{cases}$$

$$a = \frac{2^{2k+1}}{2^{2k} - 1}\left(2 + \frac{\nu}{k}\right). \tag{62}$$

Direct calculation shows that (62) is a feasible solution to (61). (Since $\nu \geqslant k(2^k - 1)$, the variables a_c are all nonnegative.)

Case 2. N even

Pick $h = (h_1, \cdots, h_N)$, where h_i, $i = 1, \cdots, N$, is a binary k tuple, $h_2 = h_5 = h_7 = h_9 = \cdots = h_{N-1} = 0$, and $h_1, h_3, h_4, h_6, h_8, \cdots, h_N$ are nonzero. Then $\gamma(h) = (N + 2)/2$, $\beta(h) = 1$ and, by (57) and (58),

$$\phi_N(h) = 2^k(N + 2) \quad \text{and} \quad \phi_{N+1}(h) = 2^{k+1}[(2^k - 1)(N + 2) - 2^k].$$

Set

$$a_c = \begin{cases} \dfrac{(2^k - 1)N - 2}{(2^k + 1)(N - 2)}, & \text{if} \quad c = f \\ 1 - a_f = \dfrac{2(N - 2^k)}{(2^k + 1)(N - 2)}, & \text{if} \quad c = g \\ 0, & \text{otherwise,} \end{cases}$$

$$a = \frac{2^{2k+1}}{2^{2k} - 1}\left(2 + \frac{\nu}{k}\right). \tag{63}$$

Direct calculation shows that (63) is a feasible solution to (61). (Again since $\nu \geqslant k(2^k - 1)$, the variables a_c are all nonnegative.)

We have now shown that (60) is an optimal solution to (59).

APPENDIX B

In this appendix we extend Theorems 1, 2, and 3 to the case when k does not divide ν. Setting $\nu = (N - 1)k + l$, where $0 \leqslant l < k$, we have $N = \lfloor(\nu + k)/k\rfloor$ where $\lfloor y \rfloor$ denotes the integer part of y.

Encoder states are labelled with binary ν tuples in the way described in Section II. Edges of the trellis are labelled with real numbers $x(s)$, where s is a binary $\nu + k$ tuple. The group $G_{k,\nu}$ is defined in the way

described in Section II; for each binary $\nu + k$ tuple t, we define a permutation of the edge labels $x(s)$ by the rule

$$g_t(x(s)) = x(s + t).$$

The symmetry g_t^* is the sequence

$$g_t^* = (g_{t^0}, g_{t^1}, g_{t^2}, \cdots),$$

where $t^0 = t$ and t^i is obtained from t^{i-1} by cycling the entries k bits to the right and moving the last k bits to the front. When $k = 2$ and $\nu = 3$,

$$g_{11001}^* = (g_{11001}, g_{01110}, g_{10011}, g_{11100}, g_{00111}, g_{11001}, \cdots).$$

In general, $t^i = t^{d+i}$ where $d = (k + \nu)/gcd(k, k + \nu)$. Given any component of the trellis and any pair of edges $x(s)$, $x(t)$ in that component there is a unique element of $\mathbf{G}_{k,\nu}^*$ interchanging $x(s)$ and $x(t)$. The proof of Theorem 1 goes through without change and we have

$$\frac{d_{\min}^2}{P} \leqslant 4N_0, \tag{64}$$

where N_0 is the minimal length of an error event.

To see that $N_0 = N$, consider an error event E with initial state (time $t = 0$) $a = (a_1 \cdots a_{N-1}a_N)$, where $a_1, \cdots, a_{N-1}$, are k tuples and a_N is an l tuple. If k tuples b_1, b_1^* are input at time 0 then at time 1 the two paths occupy states $(b_1, a_1, \cdots, a_{N-2}, \tilde{a}_{N-1})$ and $(b_1^*, a_1, \cdots, a_{N-1}, \tilde{a}_{N-1})$, where $\tilde{s}$ denotes the l tuple $(s_1 \cdots s_l)$ obtained from the k tuple $(s_1 \cdots s_k)$ by deleting the last $k - l$ bits. At time 1 the k-tuple z_Nc is input to both paths, where c is a fixed but arbitrary $k - l$ tuple. At time 2 the two paths occupy states $(z_Nc, b_1, a_1, \cdots, a_{N-3}, \tilde{a}_{N-2})$ and $(z_Nc, b_1^*, a_1, \cdots, a_{N-3}, \tilde{a}_{N-2})$. At time N, after inputs $z_{N-1}, \cdots, z_2$, the two paths occupy states $(z_2, z_3, \cdots, z_{N-1}, z_Nc, \tilde{b}_1)$ and $(z_2, z_3, \cdots, z_{N-1}, z_Nc, \tilde{b}_1^*)$. If $\tilde{b}_1 = \tilde{b}_1^*$ then the two paths remerge at time N in state $z = (z_2, z_3, \cdots, z_{N-1}, z_Nc, \tilde{b}_1)$. We denote this error event by $E(a, z; b_1, b_1^*)$. Thus by (64)

$$\frac{d_{\min}^2}{P} \leqslant 4\left(1 + \left[\frac{\nu}{k}\right]\right) \tag{65}$$

for general k and ν.

Let $S(a, z; b_1, b_1^*)$ be the orbit of $\mathbf{G}_{k,\nu}^*$ containing the error event $E(a, z; b_1, b_1^*)$. We calculate the contribution to $Q(S(a, z; b_1, b_1^*))$ made by pairs of edges in component 0 in the same way as Lemma 8. Setting

$f = b_1 + b_1^*$ this distance contribution is

$$\frac{1}{2^{\nu+k}} \sum_t [g_t(x(b_1 a_1 \cdots a_{N-1} a_N)) - g_t(x(b_1^* a_1 \cdots a_{N-1} a_N))]^2$$

$$= \frac{1}{2^{\nu+k}} \underline{x}^T [2I_{2^{\nu+k}} - 2M(f0 \cdots 0)] \underline{x}.$$

Similarly, the distance contribution made by pairs of edges in component $N-1$ is

$$\frac{1}{2^{\nu+k}} \underline{x}^T [2I_{2^{\nu+k}} - 2M((f0 \cdots 0)^{N-1})] \underline{x}.$$

Note that the first l bits of f are zero and that the last l bits of $(f0 \cdots 0)^i$ are zero for $0 \leq i \leq N-1$. Arguing as in Lemma 8 we obtain

$$2^{\nu+k} Q(S(a, z; b_1, b_1^*)) = 2NI_{2^{\nu+k}} - 2 \sum_{i=0}^{N-1} M((f0 \cdots 0)^i).$$

There are $(2^{k-l} - l)$ k tuples f for wbicb $f \neq 0$ and $\tilde{f} = 0$. Hence

$$2^{\nu+k}(2^{k-l} - 1) Q(S^N) = 2(2^{k-l} - 1) NI_{2^{\nu+k}} - 2 \sum_{\substack{f \neq 0 \\ \tilde{f} = 0}} \sum_{i=0}^{N-1} M((f0 \cdots 0)^i).$$

Setting $Q = 2^{\nu+k}(2^{k-l} - 1) Q(S^N)$ we obtain

$$\frac{d_{\min}^2}{P} \leq \frac{1}{2^{k-l} - 1} \lambda_1(Q), \tag{66}$$

which reduces to (39) when $l = 0$. The proof of Theorem 2 goes through (change "$c_i = 0 (\neq 0)$" to "the last $k - l$ digits of c_i are zero (nonzero)")

$$\lambda_1(Q) = 2^{k-l+1} N. \tag{67}$$

By (66) and (67)

$$\frac{d_{\min}^2}{P} \leq \frac{2^{k-l+1}}{2^{k-l} - 1} \left(1 + \left\lfloor \frac{\nu}{k} \right\rfloor \right) \tag{68}$$

for general k and ν.

Finally we consider the set S' of all error events of lengtb $N + 1$ for which the k tuples b_1, b_1^* input at time 0 satisfy $\tilde{b}_1 = \tilde{b}_1^* = 0$ and for which the k tuples b_2, b_2^* input at time 1 satisfy $\tilde{b}_2 = \tilde{b}_2^* = 0$. Let $E \in S'$ with initial state $a = (a_1, \cdots, a_{N-1}, a_N)$ and final state $z = (z_1, \cdots, z_{N-1}, z_N)$, where a_i, z_i, $i = 1 \cdots, N-1$ are k tuples and a_N, z_N are l tuples. At time 2 the two paths occupy states $((z_N 0 \cdots 0) + b_2,$

$b_1, a_1, \cdots, a_{N-3}, \tilde{a}_{N-2})$ and $((z_N 0 \cdots 0) + b_2^*, b_1^*, a_1, \cdots, a_{N-3}, \tilde{a}_{N-2})$. At time N the two paths occupy states $(z_2, z_3, \cdots, z_{N-1}, (z_N 0 \cdots 0) + b_2, \tilde{b}_1)$ and $(z_2, z_3, \cdots, z_{N-1}(z_N 0 \cdots 0) + b_2^*, b_1^*)$. Set $f = b_2 + b_2^*$ and $g = b_1 + b_1^*$ and define $M_i(fg)$ $i = 0, \cdots N$ as in (43). Arguing as in Lemma 9 we obtain

$$2^{\nu+k}(2^{k-l})(2^{k-l} - 1)Q(S')$$

$$= 2(2^{k-l} - 1)(2^{k-l} - 1)(N + 1)I_{2^{\nu+k}} - 2 \sum_{\substack{f \neq 0 \\ \tilde{f}=0}} \sum_{\substack{g \neq 0 \\ \tilde{g}=0}} \sum_{i=0}^{N} M_i(fg0 \cdots 0).$$

Let $\delta = 1/(2^{2(k-l)} - 1)$ and let $\bar{Q} = 2(2^{k-l} - 1)\delta Q(S^N) + (2^{k-l} - 1)^2 \delta Q(S')$. Then $2(2^{k-l} - 1)\delta + (2^{k-l} - 1)^2\delta = 1$ and so

$$\frac{d_{\min}^2}{P} \leqslant 2^{\nu+k}\lambda_1(\bar{Q}).$$

The proof of Theorem 3 goes through [change "$c_i = 0$ ($\neq 0$)" to "the last $k - l$ digits of c_i are zero (nonzero)"] and we obtain

$$\frac{d_{\min}^2}{P} \leqslant \frac{2^{2(k-l)+1}}{2^{2(k-l)} - 1}\left(2 + \left\lfloor \frac{\nu}{k} \right\rfloor\right), \tag{69}$$

for general k and ν.

AUTHORS

Helene Shapiro, B.A. (Mathematics), 1975, Kenyon College; Ph.D. (Mathematics), 1979, California Institute of Technology; Assistant Professor of Mathematics, University of Wisconsin-Madison, 1979–80; Assistant Professor of Mathematics, Swarthmore College, 1980—. Ms. Shapiro's research has been in matrix theory and deals with generalized commutativity prohlems.

Robert Calderbank, B.Sc. (Mathematics), 1975, University of Warwick; M.Sc. (Mathematics), 1976, University of Oxford; Ph.D. (Mathematics), 1980, California Institute of Technology; Bell Lahoratories, 1980—. Mr. Calderhank is presently employed in the Mathematics and Statistics Research Center.

James E. Mazo, B.S. (Physics), 1958, Massachusetts Institute of Technology; M.S. (Physics), 1960, and Ph.D. (Physics), 1963, Syracuse University; Research Associate, Department of Physics, University of Indiana, 1963–1964; Bell Lahoratories, 1964—. At the University of Indiana, Mr. Mazo worked on studies of scattering theory. At Bell Lahoratories, he has heen concerned with prohlems in data transmission and is now working in the Mathematics and Statistics Research Center. Member, IEEE.